

Roll No.

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Total No. of Pages : 02

Total No. of Questions : 09

B.Tech. (AI & ML) (Sem.-6)

BIG DATA ANALYTICS

Subject Code : BTDS-603-20

M.Code : 93675

Date of Examination : 14-06-2025

Time : 3 Hrs.

Max. Marks : 60

INSTRUCTIONS TO CANDIDATES :

1. SECTION-A is COMPULSORY consisting of TEN questions carrying TWO marks each.
2. SECTION-B contains FIVE questions carrying FIVE marks each and students have to attempt any FOUR questions.
3. SECTION-C contains THREE questions carrying TEN marks each and students have to attempt any TWO questions.

SECTION-A

1. Write briefly :

- a) List out the best practices of big data analytics.
- b) Name some final deliverables in big data.
- c) Define Big Data Analytics.
- d) What is a symbol map?
- e) Differentiate between classification and clustering.
- f) What is web usage mining?
- g) What are the measures of similarity and dissimilarity?
- h) What is data preprocessing?
- i) What is a text mining?
- j) List some big data platforms.

SECTION-B

- Write down the characteristics of big data analytics.
- Explain in detail about HDFS.
- Briefly discuss about MapReduce and YARN.
- The distance between some Indian cities are given below, apply the clustering algorithm to make three clusters. Indicate intermediate steps.

	Bathinda	Patiala	Delhi	Amritsar	Mathura
Bathinda	0	190	400	250	460
Patiala	190	0	240	225	300
Delhi	400	240	0	450	60
Amritsar	250	225	450	0	510
Mathura	460	300	60	510	0

- Write in detail the steps in operationalizing analytics model.

SECTION-C

- Write short notes on :

- Naïve Bayes
- Association Rule Mining.

- Map Reduce Framework.
 - What is web content mining? Write its different steps.

- Explain :

- Core deliverables in big data.
- Big data visualization.

NOTE : Disclosure of Identity by writing Mobile No. or Making of passing request on any page of Answer Sheet will lead to UMC against the Student.